A NOVEL GENE FOR CONTROLLING LEAF SHAPES

BACKGROUND OF THE INVENTION

1. FIELD OF THE INVENTION:

The present invention relates to a novel gene. In particular, the present invention relates to a novel gene in plants which encodes a protein having the function of controlling leaf shapes.

2. DESCRIPTION OF THE RELATED ART:

Transposons are mutagenic genes which are known to be ubiquitous in animal, yeast, bacterial, and plant genomes. Transposons are classified into two classes, Class I and Class II, depending on their transposition mechanisms. Transposons belonging to Class II are transposed in the form of DNAs without being replicated. Known Class II transposons include the Ac/Ds, Spm/dSpm and Mu elements of Zea mays (Fedoroff, 1989, Cell 56, 181-191; Fedoroff et al., 1983, Cell 35, 235-242; Schiefelbein et al., 1985, Proc. Natl. Acad. Sci. USA 82, 4783-4787), and the Tam element of Antirrhinum majus (Bonas et al., 1984, EMBO J., 3, 1015-1019). Class II transposons are widely used for gene isolation techniques which utilize transposon tagging. Such techniques utilize the fact that a transposon induces physiological and morphological changes when inserted into genes. The affected gene can be isolated by detecting such changes (Bancroft et al., 1993, The Plant Cell, 5, 631-638; Colasanti et al., 1998, Cell, 93, 593-603; Gray et al., 1997, Cell, 89, 25-31; Keddie et al., 1998, The Plant Cell, 10, 877-887; Whitham et al., 1994, Cell, 78, 1101-1115).

Transposons belonging to Class I, also referred to as retrotransposons, are replicated and transposed via RNA

intermediates. Class I transposons were first identified and characterized in Drosophila and in yeasts. However, recent studies have revealed that Class I transposons are ubiquitous in plant genomes and account for a substantial portion of the genomes (Bennetzen, 1996, Trends Microbiolo., 4, 347-353; Voytas, 1996, Science, 274, 737-738). A large majority of retrotransposons appear to be inactive. Recent studies indicate that some of these retrotransposons are activated under stress conditions such as injuries, pathogenic attacks, or cell culture (Grandbastien, 1998, Trends in Plant Science, 3, 181-187; Wessler, 1996, Curr. Biol. 6, 959-961; Wessler et al., 1995, Curr. Opin. Genet. Devel. 5, 814-821). Activation under stress conditions has been reported for Tnt1A and Tto1 in tobacco (Pouteau et al., 1994, Plant J., 5, 535-542; Takeda et al., 1988, Plant Mol. Biol., 36, 365-376), and Tos17 in rice (Hirochika et al., 1996, Proc. Natl. Acad. Sci. USA, 93, 7783-7788), for example.

The Tos17 retrotransposon of rice is one of the most-extensively studied plant Class I elements in plants. Tos17 was cloned by an RT-PCR method using a degenerate primer prepared based on a conservative amino acid sequence in reverse transcription enzyme domains between Tyl-copia retroelements (Hirochika et al., 1992, Mol. Gen. Genet., 233, 209-216). Tos17 is 4.3kb long, and has two 138 bp LTRs (long chain terminal repetitions) and PBS (primer binding sites) complementary to the 3' end of the start methionine tRNA (Hirochika et al., 1996, supra). Tos17 transcripti n is strongly activated through tissue culture, and its copy number increases with culture time. In Nipponbare, a model Japonica cultivar used for genome analysis, two copies of Tos17 are initially present, which are increased to 5 to

30 copies in a regenerated plant after tissue culture (Hirochika et al., 1996, supra). Unlike Class II transposons which were characterized in yeasts and Drosophila, Tos17 is transposed in chromosomes in random manners and causes stable mutation, and therefore provides a powerful tool for functional analysis of rice genes (Hirochika, 1997, Plant Mol. Biol. 35, 231-240; 1999, Molecular Biology of Rice (ed. by K. Shimamoto, Springer-Verlag, 43-58).

SUMMARY OF THE INVENTION

The present invention relates to a polynucleotide encoding a plant gene capable of controlling leaf shapes, the polynucleotide encoding an amino acid sequence from M t at position 1 to Val at position 690 of SEQ ID NO: 2 in the SEQUENCE LISTING, including any polynucleotide encoding an amino acid sequence in which one or more amino acids are deleted, substituted or added to the amino acid sequence.

In one embodiment of the invention, the polynucleotide may be derived from rice.

In another embodiment of the invention, the polynucleotide may be as represented by SEQ ID NO: 1 in the SEQUENCE LISTING.

The present invention further relates to methods for controlling leaf shapes in plants.

The inventors diligently conducted systematic analyses of phenotypes of plants having a newly transposed To17 copy and sequences adjoining Tos17 target sites with

respect to rice. As a result, the inventors found a narrow-leaf rice mutation obtained from Tos17 insertion, and isolated the gene responsible for this mutation by utilizing Tos17 as a tag, thereby accomplishing the present invention.

Thus, the invention described herein makes possibl the advantage of: providing a novel plant gene which can be provided by using Tos17.

This and other advantages of the present invention will become apparent to those skilled in the art upon reading and understanding the following detailed description with reference to the accompanying figures.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a photograph showing a Tos17-inserted narrow-leaf mutant rice plant (left) and a wild-type rice plant (right).

Figure 2 shows a Southern analysis autoradiogram of DNA extracted from self-crossed progeny from a narrow-leaf mutant NC0608 strain (R2 generation) and DNA extracted from a wild-type rice. On the left is shown a autoradiogram of a Southern analysis performed by using Tos17 as a probe. On the right is shown an autoradiogram of a Southern analysis performed by subcloning NC0608_0_102, which is one of the adjoining sequences of Tos17, and using it as a probe. The lane indicated as M is a lane of a λ/HindIII marker. The lane indicated as C is a control lane in which DNA obtained from a wild-type plant (Nipponbare) was electrophoresed. The lane indicated as mt is a lane in which DNA obtained

from a narrow-leaf mutant was electrophoresed.

Figure 3 is a schematic representation of a gene which control leaf shapes. Blank boxes in the figure represent introns, whereas black boxes represent exons. The downward arrow on the right-hand side of the figure represent a position at which Tos17 was inserted. The two small downward arrows near the 5' end and the 3' end represent a start codon site and a stop codon site, respectively.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention provides a novel plant gene which can be provided by using Tos17, a vector containing the same, a plant which is transformed by the novel gene, and a method of producing an improved plant including a st p of transforming a plant with the novel gene.

According to the present invention, there is provided a polynucleotide encoding a plant gene capable of controlling leaf shapes. As used herein, the term "controlling leaf shapes" means the ability to alter the leaf length and/or leaf width of a plant, thereby enhancing photosynthesis ability or imparting resistance against lodging, etc. The term "plants" encompasses both monocotyledons and dicotyledons.

A polynucleotide encoding a plant gene capable of controlling leaf shapes according to the present invention is, for example, a polynucleotide encoding an amino acid sequence from Met at position 1 to Val at position 690 of SEQ ID NO: 2 in the SEQUENCE LISTING, including any polynucleotide encoding an amino acid sequence in which one

or more amino acids are deleted, substituted or added to the aforementioned amino acid sequence.

A polynucleotide encoding a plant gene capable of controlling leaf shapes encompasses any polynucleotides which have at least about 80% sequence homology, preferably at least about 85% sequence homology, and more preferably at least about 90% sequence homology, still more preferably at least about 95% sequence homology, and most preferably at least about 99% sequence homology, with an amino acid sequence from Met at position 1 to Val at position 690 of SEQ ID NO: 2 in the SEQUENCE LISTING, so long as they are capable of controlling leaf shapes in plants. The term "sequence homology" indicates a degree of identicalness between two polynucleotide sequences to be compared with each other. The rate (%) of sequence homology between two polynucleotide sequences for comparison is calculated by, after optimally aligning the two polynucleotide sequences for comparison, obtaining a matched position number indicating the number of positions at which identical, or "matched", nucleic acid bases (e.g., A, T, C, G, U, or I) are present in both sequences, dividing the matched position number by total number of bases in the polynucleotide sequences for comparison, and multiplying the quotient by 100. The sequence homology can be calculated by using the following sequencing tools, for example: a Unix base program designated GCG Wisconsin Package (Program Manual for the Wisconsin Package, Version 8, September 1994, Genetics Computer Group, 575 Science Drive Madison, Wisconsin, USA 53711; Rice, P. (1996) Program Manual for EGCG Package, Peter Rice, The Sanger Centre, Hinxton Hall, Cambridge, CB10 1RQ, England), and the ExPASy World Wide Web molecular biology server (Geneva University Hospital and University of Geneva,

Geneva, Switzerland).

The term "control sequence" as used herein refers to a DNA sequence including a functional promoter and any related transcription elements (e.g., an enhancer, CCAAT box, TATA box, SPI site, etc.).

The term "operably linked" as used herein refers to a manner of linking a polynucleotide such that various regulation elements such as a promoter, enhancer, etc., which regulate its expression can operate within a host cell.

It is well-known to those skilled in the art that the type and kinds of control sequences may vary depending on the host cell. For example, CaMV35S promoter, nopaline synthase promoter, and the like are well-known to those skilled in the art. Any methods that are known to those skilled in the art may be used for introducing the gene into a plant body. For example, methods which utilize agrobacterium and methods which directly introduce a gene in a cell are well known. As for methods which utilize agrobacterium, the method of Nagel et al. (Microbiol. Lett. 67, 325 (1990)) may be used, for example. This method involves first transforming agrobacterium with an expression vector via electroporation, and then introducing the transformed agrobacterium into a plant cell by following a method described in Plant Molecular Biology Manual (S.B. Gelvin et al., Academic Press Publishers). Electroporation techniques and partile gun techniques are known as methods for directly introducing a gene into a cell.

Cells into which genes have been introduced are first selected based on drug resistance, e.g., hygromycin

resistance, and then regenerated into plant bodies by using usual methods.

The terminology and laboratory procedures described throughout the present specification are directed to those which are well-known and commonly employed in the art. Standard techniques may be used for recombination methods, polynucleotide synthesis, microorganisms culturing, and transformation (e.g., electroporation). Such techniques and procedures are generally known from various standard textbooks available in the field or by way of the present specification (including a generally-referenced textbook by Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd edition (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.). Such literature is incorporated herein by reference.

The polynucleotide according to the present invention can be obtained by using the method described herein, for example. However, the polynucleotide according to the present invention may also be obtained by any chemical synthesis process based on the sequence disclosed her in. For example, the polynucleotide according to the present invention may be synthesized by using a polynucleotide synthesizer available from Applied Bio Systems in accordance with the instructions provided by the manufacturer.

Methods of PCR amplification are well-known in the art (PCR Technology: Principles and Applications for DNA Amplification, ed. HA Erlich, Freeman Press, NewYork, NY (1992); PCR Protocols: A Guide to Methods and Applications, Innis, Gelfland, Snisky, and White, Academic Press, San Diego, CA(1990); Mattila et al. (1991) Nucleic Acids REs.

19: 4967; Eckert, K.A. and Kunkel, T. A. (1991) PCR Methods and Applicatdions 1: 17; PCR, McPherson, Quirkes, and Taylor, IRL Press, Oxford). Such literature is incorporated herein by reference.

(Examples)

Hereinafter, the present invention will be described by way of examples which are of illustrative but not limitative nature.

(Example 1: Activation of Tos17 via culture

Using fully ripened seeds of Nipponbare, which is a variety of Japonica subspecies, induction of calli and cell suspension culture were carried out as described earlier (Hirochika et al., 1996, supra). The activation of Tos17 was carried out following the method of Ohtsuki (1990) (rice protoplast culture system, Food and Agricultural Research Development Association). In summary, fully ripened seeds of rice were cultured in an MS medium having 2,4-dichlorophenoxyacetic acid (2,4-D) added thereto (2 mg/ml) (Ohtsuki (1990), supra) (25℃, 1 month), to induce callus formation. The resultant calluses were cultured for 5 months in an N6 liquid medium having 2,4-D added thereto (Ohtsuki (1990), supra), and thereafter placed on a redifferentiation medium (Ohtsuki (1990), supra), wher by redifferentiated rice plants were obtained (first generation (R1) plants).

(Example 2: Isolation and identification of narrow-leaf mutants

Utilizing each of the regenerated R1 rice plants obtained according to Example 1, about 1000 R1 seeds were collected from each strain and sown on a paddy field to obtain

second generation (R2) plants, which were subjected to a morphological analysis. As a result of observing the phenotypes of the respective plant bodies in the R2 group, it was learned that about 1/4 of the R2 group of the NC0608 strain exhibit the "narrow-leaf" phenotype (Figure 1). In the paddy field, the Tos17-inserted narrow-leaf mutants had their leaf length reduced to about 90% in the flag leaf and all leaves down to the third leaf therefrom; and they also had their leaf width reduced to about 78%, about 70%, about 71%, about 69%, respectively, in the flag leaf and all leav s down to the third leaf therefrom (Figure 1, left), as compared with the wild type (Figure 1, right). This suggested that the narrow-leaf phenotype of NC0608 is caused by recessive mutation at a single gene locus.

(Example 3: Isolation of causative gene for narrow-leaf mutations)

In order to identify and isolate the causative gene for narrow-leaf mutations from the NC0608 strain obtained according to Example 2, linkage analysis with respect to the Tos17 gene was performed on a group part of which was separable as narrow-leaf mutations. In order to show that recessive mutation at a single gene locus is responsible for the mutations, adjoining portions of a target site (Ts) of the NC0608 strain at which Tos17 had been transpos - inserted were amplified first.

From the group of R2 rice plants(self-cross d progeny from the NC0608 strain) obtained according to Example 2, individuals exhibiting mutation were identified from normal individuals. DNA was prepared from both groups of individuals by using a CTAB method (Murray and Thompson, 1980, Nucleic Acids Res. 8, 4321-4325). The DNA obtained

from individuals exhibiting narrow-leaf mutation and the DNA obtained from normal individuals were each digested with restriction enzyme XbaI, and after agarose electrophoresis, were allowed to adsorb to nylon membranes. DNA fragments which were obtained from Tos17 through digestion by XbaI and BamHI were labeled with ^{32}P-dCTP. By using these as probes, a Southern hybridization was performed (Figure 2, left). As seen from the Southern analysis autoradiogram shown on the left-hand side in Figure 2, it was learned the Tos17 band (about 6600 bp) indicated by an arrow was observed in narrow-leaf mutations as a homozygous band, but not in normal individuals, and that the Tos17 band indicated by the arrow was completely linked with the narrow-leaf mutation phenotype. From these results, it was conclud d that the DNA which is represented by the band which hybridizes to the Tos17 probe indicated by the arrow contains a causative gene, such that Tos17, when inserted in a genome region represented by this band, generates narrow-leaf mutations as the genotype becomes homozygous. Accordingly, a portion of the causative gene for the narrow-leaf mutations, i.e., a sequence adjoining Tos17, was isolated through TAIL-PCR reactions using this DNA as a template. The amplification of the Tos17 target site sequence was accomplished by TAIL-PCR employing the total DNA (Liu Y-G. et al., 1995, Genomics, 25, 674-681, Liu Y-G. et al., 1995, Plant J., 8, 457-463). In summary, by using as a template the total DNA from a regenerated plant having a new Tos17 target site, three TAIL-PCR amplification reactions were performed, using the following three sets of primers: (1st reaction) Tos17 Tail3, GAGAGCATCATCGGTTACATCTTCTC and AD1 (arbitrarily degenerated primer 1) NGTCGA (G/C) (A/T) GANA (A/T) GAA; (2nd reaction) Tos17 Tail4, ATCCACCTTGAGTTTGAAGGG and AD1;

and (3rd reaction) Tos17 Tail5, CATCGGATGTCCAGTCCATTG and AD1. Next, the respective TAIL-PCR products were subjected to an agarose electrophoresis and then a simple column purification. By directly applying them to a sequencer (Model 377 available from ABI), sequencing was performed.

Four new target sites (Ts) for Tos17 insertion were identified as a result of sequencing the adjoining sequences of Tos17 in the NC0608 strain.

Next, a Southern analysis was performed by subcloning NC0608_0_102, one of the adjoining sequences of Tos17, and using it as a probe. The results are shown on the right-hand side in Figure 2. As seen from the autoradiogram on the right-hand side in Figure 2, the Tos17-adjoining sequence NC0608_0_102 hybridized to the DNA fragment located at the same position as that indicated in the Southern analysis in which Tos17 was used as a probe. The results were consistent for all of the 62 strains that were examined. This indicates that the subclone NC0608_0_102 contains a portion of the causative gene for the narrow-leaf mutation, and that NC0608_0_102 is an adjoining sequence of the causative gene for the narrow-leaf mutation.

(Example 4: Structural analysis of the causative gene for narrow-leaf mutation)

Relying on the adjoining sequence obtained according to Example 3, the inventors attempted to determine the complete structure of cDNA which was transcribed from the gene containing the adjoining sequence NC0608_0_102 through a PCR screening using a cDNA library and Cap Site cDNA (Nippongene). By using the wild-type (Nipponbare) DNA

as a template, the inventors attempted to determine the complete structure of the genomic DNA of the gene containing NC0608_0_102 through a PCR using a primer which is design d from the cDNA and through the aforementioned TAIL-PCR.

The cDNA library was previously prepared in the laboratory of the inventors. The method of preparation can be summarized as follows. First, by using an ISOGEN solution (Nippongene), the total RNA was extracted from a callus of a wild-type rice plant which had been cultured in the aforementioned MS medium. By using an oligo(dT)cellulose column contained in an mRNA purification kit (Stratagen), poly(A)mRNA was obtained from the total RNA. Following usual methods, cDNA was synthesized from the resultant poly(A)mRNA. Thus, a cDNA library was constructed in a Hybri ZAP-II vector (Stratagene).

The cDNA and genomic DNA of the gene containing the adjoining sequence NC0608_0_102 were partially amplified through the below-described four-step PCR reactions and three-step PCR reactions, respectively. All of the amplified fragments were sequenced by using a 377 sequenc r (Perkin Elmer) for both directions.

(cDNA)

First step: Using cDNA library as a template, a PCR reaction was carried out by using a pair of primers specific to the adjoining sequence NC0608_0_102 to confirm that a portion of this adjoining sequence is contained in the cDNA library: NC0608_0_102F ACGGAGACACCTCGTAAACC and NC0608_0_102R1 AAGGCCGACTATTGTTGACC.

Second step: Using the cDNA library as a template, a PCR reaction was carried out by using NC0608_0_102F and Hybri ZAP B (Stratagene), which is a primer specific to Hybri ZAP-II vector. Thus, a fragment which partially overlaps with NC0608_0_102 and which contains the 3' region of cDNA along with the poly(A) binding site was obtained.

Third step: Using the cDNA library as a templat , a PCR reaction was carried out by using Hybri ZAP A (Stratagene), which is a primer specific to Hybri ZAP-II vector, and NC0608_0_102R2 CCTGCAATGTTACCTCTGGC, which is a primer specific to NC0608_0_102. Thus, a 5' fragment which partially overlaps with NC0608_0_102 was obtained.

Fourth step: Using Cap Site cDNA (Nipponegne) as a template, a PCR reaction was carried out by using 1RC2 (Nippongene), which is a primer specific to Cap Site, and TGACAGGTCAGACTGATCAACCGG, which is a primer specific to the fragment obtained in the third step. Thus, a fragment which partially overlaps with the fragment obtained in the third step and which contains the 5' region of cDNA along with the transcription start point (cap site).

(Genomic DNA)

First step: Using the total DNA of Nipponbare, two reactions of TAIL-PCR were carried out using the following two sets of primers to obtain a 5' fragment which partially overlaps with the NC0608_0_102: (first reaction: NC0608_0_102R2 and AD1 employed in Example 3; second reaction: NC0608_0_102R3 TAGGCAATCCGGCAATGTCC and AD1)

Second step: Using the total DNA of Nipponbare, a

PCR reaction was carried out using a primer (CTAGAAGCAAAATCTTGAAGCTGC) which is specific to the fragment obtained in the first step and a primer (AGTGTTCTTCGCACCTCGCG) which is specific to the cDNA fragment obtained in the fourth step PCR. Thus, a 5' fragment which partially overlaps with the fragment obtained in the first step was obtained.

Third step: Using the total DNA of Nipponbare, a PCR reaction was carried out using a primer (TGCCTCGCCCTCGGCGATGG) which is specific to the fragment obtained in the second step and a primer (AATATTTCAAATCACACTAC) which is specific to the 5' region of the cDNA fragment obtained in the fourth step PCR. Thus, a 5' fragment which partially overlaps with the fragment obtained in the second step was obtained.

The cDNA and genomic DNA structures of the narrow-leaf gene are shown together in Figure 3. This gene has 11 introns and encodes 690 amino acids, and yet finds no similar genes registered in existing databases. Thus, it was confirmed that this gene is novel. It was learned that Tos17 had been inserted between the 9th and the 10th bases from the 5' end of the 12th exon region. An amino acid sequence encoded by this gene showed very high homology with a gene in Arabidopsis thaliana having an unknown function.

The above examples are illustrative, and by no means limitative, of various aspects of the present invention and the manners in which the oligonucleotide according to the present invention can be made and utilized.

Thus, according to the present invention, a n vel

polynucleotide is provided which is capable of controlling leaf shapes, the polynucleotide being of use in plant breeding. By introducing the present polynucleotide into plants and artificially controlling leaf shapes, it is expected that enhancement of photosynthesis ability or provision of resistance against lodging, etc., can be attained.

Various other modifications will be apparent to and can be readily made by those skilled in the art without departing from the scope and spirit of this invention. Accordingly, it is not intended that the scope of the claims appended hereto be limited to the description as set forth herein, but rather that the claims be broadly construed.

SEQUENCE LISTING

<110> National Institute of Agrobiological Resources, Ministry of Agriculture, Forestry and Fisheries
Bio-Oriented Technology Research Advancement Institution Society for Techno-Innovation of Agriculture, Forestry, and Fisheries

<120> A novel gene which controls leaf shape and size

<130>

<140> JP
<141> 2000-03-23

<160> 3

<170> PatentIn Ver. 2.1

<210> 1
<211> 2468
<212> DNA
<213> Oryza sativa

<220>
<221> CDS
<222> (198)..(2270)

<300>

<400> 1
aaaaaaatat ttcaaatcac actacactct ccgtcgtctc ctctcctctc ctctcctccc 60

cctctcctcc gcctctctcg catctgaggc tccgatcgcc ggcgacccca gccagaatcc 120

gccgccccgt ctcgccctcc ccgctcgacg agaccgcgcc gagcggcgaa gaggcctagt 180

gttcttcgca cctcgcg atg agt agc gcg gtg aag gac cag ctt cac cag 230
Met Ser Ser Ala Val Lys Asp Gln Leu His Gln
1 5 10

atg tcg acg aca tgc gat tcg ctt cta ctg gag ctc aat gtg att tgg 278
Met Ser Thr Thr Cys Asp Ser Leu Leu Leu Glu Leu Asn Val Ile Trp
15 20 25

gat gag gtc ggt gag ccc gac acg acg agg gac agg atg ctg ctg gag 326
Asp Glu Val Gly Glu Pro Asp Thr Thr Arg Asp Arg Met Leu Leu Glu
30 35 40

ctc gag cag gag tgc ctg gag gtc tac agg cgg aag gtc gac cag gcg 374
Leu Glu Gln Glu Cys Leu Glu Val Tyr Arg Arg Lys Val Asp Gln Ala
45 50 55

aac cgg agc cgc gcc cag ctg cgg aag gcc atc gcc gag ggc gag gca 422
Asn Arg Ser Arg Ala Gln Leu Arg Lys Ala Ile Ala Glu Gly Glu Ala
60 65 70 75

gag ctc gcc ggc atc tgc tca gcc atg ggc gag ccg ccc gtg cac gtt 470
Glu Leu Ala Gly Ile Cys Ser Ala Met Gly Glu Pro Pro Val His Val
80 85 90

aga cag tca aat cag aag ctt cat ggc tta aga gag gag ttg aat gca 518
Arg Gln Ser Asn Gln Lys Leu His Gly Leu Arg Glu Glu Leu Asn Ala
95 100 105

att gtt ccg tat ttg gaa gaa atg aaa aag aaa aag gtc gaa cga tgg 566
Ile Val Pro Tyr Leu Glu Glu Met Lys Lys Lys Lys Val Glu Arg Trp

```
          110                 115                 120

     aac cag ttt gtt cat gtc ata gag cag att aag aaa att tcg tct gaa   614
     Asn Gln Phe Val His Val Ile Glu Gln Ile Lys Lys Ile Ser Ser Glu
5      125                 130                 135

     ata agg cca gcc gat ttt gtt ccc ttt aaa gtt ccg gtt gat cag tct   662
     Ile Arg Pro Ala Asp Phe Val Pro Phe Lys Val Pro Val Asp Gln Ser
     140                 145                 150                 155
10
     gac ctg tca tta aga aag ctt gat gag ttg acg aag gac ctg gna tcc   710
     Asp Leu Ser Leu Arg Lys Leu Asp Glu Leu Thr Lys Asp Leu Xaa Ser
                    160                 165                 170

15   ctt cag aag gag aag agc gat cgg cta aag caa gtg ata gaa cat ttg   758
     Leu Gln Lys Glu Lys Ser Asp Arg Leu Lys Gln Val Ile Glu His Leu
                175                 180                 185

     aat tct ttg cat tcc tta tgt gag gtg ctt ggc ata gat ttc aag caa   806
20   Asn Ser Leu His Ser Leu Cys Glu Val Leu Gly Ile Asp Phe Lys Gln
          190                 195                 200

     aca gta tat gag gtg cac cct agc ttg gac gaa gct gaa gga tca aag   854
     Thr Val Tyr Glu Val His Pro Ser Leu Asp Glu Ala Glu Gly Ser Lys
25     205                 210                 215

     aac ctg agc aac act aca att gag agg ctt gct gct gcc gca aac aga   902
     Asn Leu Ser Asn Thr Thr Ile Glu Arg Leu Ala Ala Ala Ala Asn Arg
     220                 225                 230                 235
30
     ctg cgt gaa atg aag atc caa agg atg caa aag ctt caa gat ttt gct   950
     Leu Arg Glu Met Lys Ile Gln Arg Met Gln Lys Leu Gln Asp Phe Ala
                    240                 245                 250
```

tct agc atg ctc gag cta tgg aat ctc atg gat act cca ctt gaa gag 998
Ser Ser Met Leu Glu Leu Trp Asn Leu Met Asp Thr Pro Leu Glu Glu
255 260 265

cag cag atg ttt cag aat ata aca tgc aat att gct gct tca gaa caa 1046
Gln Gln Met Phe Gln Asn Ile Thr Cys Asn Ile Ala Ala Ser Glu Gln
270 275 280

gag ata act gaa cca aac acc ctc tcc aca gat ttc ctg aat tat gtc 1094
Glu Ile Thr Glu Pro Asn Thr Leu Ser Thr Asp Phe Leu Asn Tyr Val
285 290 295

gaa tct gag gtg tta agg ctt gaa caa ctg aaa gca agt aag atg aaa 1142
Glu Ser Glu Val Leu Arg Leu Glu Gln Leu Lys Ala Ser Lys Met Lys
300 305 310 315

gat ctt gtt tta aaa aag aaa gca gaa cta gaa gag cat aga aga cgt 1190
Asp Leu Val Leu Lys Lys Lys Ala Glu Leu Glu Glu His Arg Arg Arg
320 325 330

gct cat ctt gtt ggc gag gaa ggt tat gca gag gag ttt agc att gaa 1238
Ala His Leu Val Gly Glu Glu Gly Tyr Ala Glu Glu Phe Ser Ile Glu
335 340 345

gct att gaa gct gga gct att gat ccc tca cta gta ctt gaa caa att 1286
Ala Ile Glu Ala Gly Ala Ile Asp Pro Ser Leu Val Leu Glu Gln Ile
350 355 360

gaa gct cac att gca aca gtg aaa gag gaa gct ttt agc cgg aag gat 1334
Glu Ala His Ile Ala Thr Val Lys Glu Glu Ala Phe Ser Arg Lys Asp
365 370 375

```
att ctt gag aaa gtt gaa aga tgg caa aat gct tgt gaa gag gaa gcc  1382
Ile Leu Glu Lys Val Glu Arg Trp Gln Asn Ala Cys Glu Glu Glu Ala
380                 385                 390                 395

tgg ctg gaa gat tac aac aaa gat gat aat cgt tac aat gct ggg agg  1430
Trp Leu Glu Asp Tyr Asn Lys Asp Asp Asn Arg Tyr Asn Ala Gly Arg
                400                 405                 410

gga gca cat cta aca cta aag agg gct gaa aag gct cgt act ttg gtc  1478
Gly Ala His Leu Thr Leu Lys Arg Ala Glu Lys Ala Arg Thr Leu Val
            415                 420                 425

aac aag att cct gga atg gta gat gtt ttg aga aca aaa att gct gca  1526
Asn Lys Ile Pro Gly Met Val Asp Val Leu Arg Thr Lys Ile Ala Ala
        430                 435                 440

tgg aaa aat gaa cga gga aag gag gat ttc aca tat gat ggt gtt agc  1574
Trp Lys Asn Glu Arg Gly Lys Glu Asp Phe Thr Tyr Asp Gly Val Ser
    445                 450                 455

ctt tcg tca atg ctt gat gaa tat atg ttc gtt cgt cag gag aaa gag  1622
Leu Ser Ser Met Leu Asp Glu Tyr Met Phe Val Arg Gln Glu Lys Glu
460                 465                 470                 475

caa gag aag aag aga caa agg gat cag aag aag ctc cag gat cag ctc  1670
Gln Glu Lys Lys Arg Gln Arg Asp Gln Lys Lys Leu Gln Asp Gln Leu
                480                 485                 490

aaa gcg gag cag gaa gct ttg tac gga tca aaa ccc agt cca tcc aag  1718
Lys Ala Glu Gln Glu Ala Leu Tyr Gly Ser Lys Pro Ser Pro Ser Lys
            495                 500                 505

ccc cta agt aca aag aag gca cct agg cac tct atg ggt ggt gca aac  1766
```

```
Pro Leu Ser Thr Lys Lys Ala Pro Arg His Ser Met Gly Gly Ala Asn
        510                 515                 520

cga agg cta tct ctt ggt gga gcc acc atg caa ccc ccg aag act gat   1814
Arg Arg Leu Ser Leu Gly Gly Ala Thr Met Gln Pro Pro Lys Thr Asp
    525                 530                 535

ata ctg cat tca aag tct gtt cgt gct gcc aag aaa act gaa gaa atc   1862
Ile Leu His Ser Lys Ser Val Arg Ala Ala Lys Lys Thr Glu Glu Ile
540                 545                 550                 555

ggc act ttg tcc cct agt agt agt aga ggt ttg gac att gcc gga ttg   1910
Gly Thr Leu Ser Pro Ser Ser Ser Arg Gly Leu Asp Ile Ala Gly Leu
                560                 565                 570

cct atc aag aag ttg tct ttc aat gcc agt act cta cgt gag acg gag   1958
Pro Ile Lys Lys Leu Ser Phe Asn Ala Ser Thr Leu Arg Glu Thr Glu
            575                 580                 585

aca cct cgt aaa cct ttt gct cag atc aca cca gga aac agt gtc tcg   2006
Thr Pro Arg Lys Pro Phe Ala Gln Ile Thr Pro Gly Asn Ser Val Ser
        590                 595                 600

tcg acg cct gtg cgc cct atc acc aat aac act gag gat gat gag aac   2054
Ser Thr Pro Val Arg Pro Ile Thr Asn Asn Thr Glu Asp Asp Glu Asn
    605                 610                 615

agg act ccg aag aca ttt aca gca ctg aat ccc aag act ccg atg act   2102
Arg Thr Pro Lys Thr Phe Thr Ala Leu Asn Pro Lys Thr Pro Met Thr
620                 625                 630                 635

gtt acg gct cca atg cag atg gca atg act ccc tct ctg gcc aac aag   2150
Val Thr Ala Pro Met Gln Met Ala Met Thr Pro Ser Leu Ala Asn Lys
```

640 645 650

gtt tca gca act cca gtt tcc ctt gtt tac gac aag cca gag gta aca 2198
Val Ser Ala Thr Pro Val Ser Leu Val Tyr Asp Lys Pro Glu Val Thr
655 660 665

ttg cag gag gac atc gac tac tcc ttt gaa gaa agg cgg ctc gcc atc 2246
Leu Gln Glu Asp Ile Asp Tyr Ser Phe Glu Glu Arg Arg Leu Ala Ile
670 675 680

tat ctg gcc agg caa atg gtt taa ctgttgatca atttatgtac gtagttgaaa 2300
Tyr Leu Ala Arg Gln Met Val
685 690

tctgactgca ttttcttgtc ggtggccatt gcgtatgttg gtcaacaata gtcggccttt 2360

ccagtagcac tattctgatt tactgcaatt gttttaatgt ttttacaac cagtaaaaca 2420

gcctctataca ttagcttgct cactaaaaaa aaaaaaaaaa aaaaaaaa 2468

<210> 2
<211> 690
<212> PRT
<213> Oryza sativa

<400> 2
Met Ser Ser Ala Val Lys Asp Gln Leu His Gln Met Ser Thr Thr Cys
1 5 10 15
Asp Ser Leu Leu Leu Glu Leu Asn Val Ile Trp Asp Glu Val Gly Glu
20 25 30
Pro Asp Thr Thr Arg Asp Arg Met Leu Leu Glu Leu Glu Gln Glu Cys
35 40 45
Leu Glu Val Tyr Arg Arg Lys Val Asp Gln Ala Asn Arg Ser Arg Ala

50 55 60

Gln Leu Arg Lys Ala Ile Ala Glu Gly Glu Ala Glu Leu Ala Gly Ile

65 70 75 80

Cys Ser Ala Met Gly Glu Pro Pro Val His Val Arg Gln Ser Asn Gln

85 90 95

Lys Leu His Gly Leu Arg Glu Glu Leu Asn Ala Ile Val Pro Tyr Leu

100 105 110

Glu Glu Met Lys Lys Lys Lys Val Glu Arg Trp Asn Gln Phe Val His

115 120 125

Val Ile Glu Gln Ile Lys Lys Ile Ser Ser Glu Ile Arg Pro Ala Asp

130 135 140

Phe Val Pro Phe Lys Val Pro Val Asp Gln Ser Asp Leu Ser Leu Arg

145 150 155 160

Lys Leu Asp Glu Leu Thr Lys Asp Leu Xaa Ser Leu Gln Lys Glu Lys

165 170 175

Ser Asp Arg Leu Lys Gln Val Ile Glu His Leu Asn Ser Leu His Ser

180 185 190

Leu Cys Glu Val Leu Gly Ile Asp Phe Lys Gln Thr Val Tyr Glu Val

195 200 205

His Pro Ser Leu Asp Glu Ala Glu Gly Ser Lys Asn Leu Ser Asn Thr

210 215 220

Thr Ile Glu Arg Leu Ala Ala Ala Ala Asn Arg Leu Arg Glu Met Lys

225 230 235 240

Ile Gln Arg Met Gln Lys Leu Gln Asp Phe Ala Ser Ser Met Leu Glu

245 250 255

Leu Trp Asn Leu Met Asp Thr Pro Leu Glu Gln Gln Gln Met Phe Gln

260 265 270

Asn Ile Thr Cys Asn Ile Ala Ala Ser Glu Gln Glu Ile Thr Glu Pro

275 280 285

Asn Thr Leu Ser Thr Asp Phe Leu Asn Tyr Val Glu Ser Glu Val Leu

290 295 300

Arg Leu Glu Gln Leu Lys Ala Ser Lys Met Lys Asp Leu Val Leu Lys

305 310 315 320

Lys Lys Ala Glu Leu Glu Glu His Arg Arg Arg Ala His Leu Val Gly
325 330 335
Glu Glu Gly Tyr Ala Glu Glu Phe Ser Ile Glu Ala Ile Glu Ala Gly
340 345 350
Ala Ile Asp Pro Ser Leu Val Leu Glu Gln Ile Glu Ala His Ile Ala
355 360 365
Thr Val Lys Glu Glu Ala Phe Ser Arg Lys Asp Ile Leu Glu Lys Val
370 375 380
Glu Arg Trp Gln Asn Ala Cys Glu Glu Glu Ala Trp Leu Glu Asp Tyr
385 390 395 400
Asn Lys Asp Asp Asn Arg Tyr Asn Ala Gly Arg Gly Ala His Leu Thr
405 410 415
Leu Lys Arg Ala Glu Lys Ala Arg Thr Leu Val Asn Lys Ile Pro Gly
420 425 430
Met Val Asp Val Leu Arg Thr Lys Ile Ala Ala Trp Lys Asn Glu Arg
435 440 445
Gly Lys Glu Asp Phe Thr Tyr Asp Gly Val Ser Leu Ser Ser Met Leu
450 455 460
Asp Glu Tyr Met Phe Val Arg Gln Glu Lys Glu Gln Glu Lys Lys Arg
465 470 475 480
Gln Arg Asp Glu Lys Lys Leu Gln Asp Glu Leu Lys Ala Glu Gln Glu
485 490 495
Ala Leu Tyr Gly Ser Lys Pro Ser Pro Ser Lys Pro Leu Ser Thr Lys
500 505 510
Lys Ala Pro Arg His Ser Met Gly Gly Ala Asn Arg Arg Leu Ser Leu
515 520 525
Gly Gly Ala Thr Met Gln Pro Pro Lys Thr Asp Ile Leu His Ser Lys
530 535 540
Ser Val Arg Ala Ala Lys Lys Thr Glu Glu Ile Gly Thr Leu Ser Pro
545 550 555 560
Ser Ser Ser Arg Gly Leu Asp Ile Ala Gly Leu Pro Ile Lys Lys Leu
565 570 575
Ser Phe Asn Ala Ser Thr Leu Arg Glu Thr Glu Thr Pro Arg Lys Pro

580 585 590
Phe Ala Gln Ile Thr Pro Gly Asn Ser Val Ser Ser Thr Pro Val Arg
595 600 605
Pro Ile Thr Asn Asn Thr Glu Asp Asp Glu Asn Arg Thr Pro Lys Thr
610 615 620
Phe Thr Ala Leu Asn Pro Lys Thr Pro Met Thr Val Thr Ala Pro Met
625 630 635 640
Gln Met Ala Met Thr Pro Ser Leu Ala Asn Lys Val Ser Ala Thr Pro
645 650 655
Val Ser Leu Val Tyr Asp Lys Pro Glu Val Thr Leu Gln Glu Asp Ile
660 665 670
Asp Tyr Ser Phe Glu Glu Arg Arg Leu Ala Ile Tyr Leu Ala Arg Gln
675 680 685
Met Val
690

<210> 3
<211> 4574
<212> DNA
<213> Oryza sativa

<400> 3
aaaaaaatat ttcaaatcac actacactct ccgtcgtctc ctctcctctc ctctcctccc 60
cctctcctcc gcctctctcg catctgaggc tccgatcgcc ggcgacccca gccagaatcc 120
gccgccccgt ctcgccctcc ccgctcgacg agaccgcgcc gagcggcgaa gaggcctagt 180
gttcttcgca cctcgcgatg agtagcgcgg tgaaggacca gcttcaccag atgtcgacga 240
catgcgattc gcttctactg gagctcaatg tatgtcaccg cttgccgatt caaccatttc 300
ccggctactc gtgttggttc tggcatggca gtggaggatt tacggggttt ttttcttctc 360
tcgttctgtt tcaggtgatt tgggatgagg tcggtgagcc cgacacgacg agggacagga 420
tgctgctgga gctcgagcag gagtgcctgg aggtctacag gcggaaggtc gaccaggcga 480
accggagccg cgcccagctg cggaaggcca tcgccgaggg cgaggcagag ctcgccggca 540
tctgctcagc catgggcgag ccgcccgtgc acgttagaca ggttagtttc tggctccacc 600

aatggctgta aaagaggtat cgcatggttg gatcaaaaga tggaagtcga attcctgtgg 660
aactgtgcta attggcgatg gaagaaaagg aagatttagt agagaactaa aagctacgat 720
ttctgttgta agatgatagt actactgctt gcattgttga tctgatggag gtaaaccgtg 780
tagaactcca tcagcagtta acatttttct aactgattag tagtagcgta tcaatatatt 840
aagggaaagt gttggcagag cttacattic tttctcactt ctatttctcga ctttatgccc 900
agttactgct caatcggttc tatacttttt actgctgttc ccatgcatta gcaatttagg 960
atatatgttt tgtaaaattt atctgtttcc ttcagtttga atatgttcag catgaataat 1020
atatttactg ttttaccggc agcatgacta agttactgcc tcaagtacgt tttatttgtt 1080
gaatacattc taccttcttg actaatcaat tctgcttgac tgtagatttt agcacttcct 1140
cagccattca tgcagtaaca tgcatttcat ctgaaatttt gcagtcaaat cagaagcttc 1200
atggcttaag agaggagttg aatgcaattg ttccgtattt ggaagaaatg aaaaagaaaa 1260
aggtcgaacg atggaaccag tttgttcatg tcatagagca gattaagaaa atttcgtctg 1320
aaataaggcc agccgatttt gttcccttta aagttccggt tgatcagtct gaccigtcat 1380
taagaaagct tgatgagttg acgaaggacc tggaatccct tcagaaggag aaggtcatca 1440
tcactaatac catctttatc catttticacc agtcatgttg tcatcgtgtc tctatctatc 1500
aagaatcctt ttcatttctt gtataaaatc tcactatgcc atatacatgt ttgtttctca 1560
cagagcgatc ggctaaagca agtgatagaa catttgaatt ctttgcattc cttatgtgag 1620
gtgcttggca tagatttcaa gcaaacagta tatgaggtgc accctagctt ggacgaagct 1680
gaaggatcaa agaacctgag caacactaca attgagaggc ttgctgctgc cgcaaacaga 1740
ctgcgtgaaa tgaagatcca aaggatgcaa aaggtcagca ttgcctgtac cattgtagag 1800
gtatcaatga acactttcag tctttaactt ggttaatctg attctggcag cttcaagatt 1860
ttgcttctag catgctcgag ctatggaatc tcatggatac tccactigaa gagcagcaga 1920
tgtttcagaa tataacatgc aatattgctg cttcagaaca agagataact gaaccaaaca 1980
ccctctccac agatttcctg aattatgtaa ttiatcatca ctgagattgc aaaaatttat 2040
gttcgtactg tgttatattt tcattaagat atgaatgttc atcgactata cttataactg 2100
taggtcgaat ctgaggtgtt aaggcttgaa caactgaaag caagtaagat gaaagatctt 2160
gttttaaaaa agaaagcaga actagaagag catagaagac gtgctcatct tgttggcgag 2220
gaaggttatg cagaggagtt tagcattgaa gctattgaag ctggtaagat actctcctgc 2280
cttactgcct tttattgtgc ctgacaagtc ataccagaca gagttcatat acctggtctg 2340
tgttctgttc gcaggagcta ttgatcccic actagtactt gaacaaattg aagctcacat 2400
tgcaacagtg aaagaggaag ctttagccg gaaggatatt cttgagaaag ttgaaagatg 2460
gcaaaatgct tgtgaagagg aagcctggct ggaagattac aacaaagtat ggatgctagc 2520
tgaagctacg tggtcttgt atattgttt agcaaataat gtggtactga tatctcctgg 2580

ctttggcttt ttttaggatg ataatcgtta caatgctggg aggggagcac atctaacact 2640
aaagagggct gaaaaggctc gtactttggt caacaagatt cctggtaatg ttactcaatg 2700
atttatgtgt ttggaacttc cttatcaagt gcatatttaa tttacaattt taactcttgc 2760
cattactaca atctgatatc ctgctgattt gtgctgagca ggaatggtag atgttttgag 2820
aacaaaaatt gctgcatgga aaaatgaacg aggaaaggag gatttcacat atgatggtgt 2880
aggttttctt actcttacac attacattga tcgggtctat ttttgtttct tgctgaagtg 2940
cctttcttgc aattcttaca ggttagcctt tcgtcaatgc ttgatgaata tatgttcgtt 3000
cgtcaggaga aagagcaaga gaagaagaga caaagggtat tatgctctcg cctaatattc 3060
atgtattgtc taaatcatct tttcaccttc tgtgaatacg ctctaatact tgaatatacc 3120
tgcaggatca gaagaagctc caggatcagc tcaaagcgga gcaggaagct ttgtacggat 3180
caaaacccag tccatccaag cccctaagta caaagaaggc acctaggcac tctatgggtg 3240
gtgcaaaccg aaggctatct cttggtggag ccaccatgca acccccgaag actgatatac 3300
tgcattcaaa gtctgttcgt gctgccaaga aaactgaaga aatcggcact ttgtccccta 3360
gtaagcccta ctagctatca tgtgtcgata tatttctttt tcctcttatt ttcacttgaa 3420
catatgtcta actcaagcaa acaatatcag gtagtagagg tttggacatt gccggattgc 3480
ctatcaagaa gttgtctttc aatgccagta ctctacgtga gacggagaca cctcgtaaac 3540
ctttgctca gatcacacca ggaaacagtg tctcgtcgac gcctgtgcgc cctatcacca 3600
ataacactga ggatgatgag aacaggactc cgaagacatt tacagcactg aatcccaaga 3660
ctccgatgac tgttacggct ccaatgcaga tggcaatgac tccctctctg gccaacaagg 3720
tttcagcaac tccagtttcc cttgtttacg acaagccaga ggtaacattg caggaggaca 3780
tcgactactc ctttgaagaa aggcggctcg ccatctatct ggccaggcaa atggtttaac 3840
tgttgatcaa tttatgtacg tagttgaaat ctgactgcat tttcttgtcg gtggccattg 3900
cgtatgttgg tcaacaatag tcggcctttc cagtagcact attctgattt actgcaattg 3960
ttttaatgtt ttctacaacc agtaaaacag ctctatacat tagcttgctc actactcagt 4020
acagctttct cggcagcacg aaacatttct gttctctttg atgaatactt cttgctgtgg 4080
atagggatag ttactgttac atatactgta tgcccttcag aatagaaacc tgttagtacg 4140
ggaggtatta taggaaggat cgttttggaa ttttggtggt tagcctgcac agtaagttcc 4200
atcagtttct ggattgtccc tcgcaaagaa aaaagttttc ttgatttctgg taatttcgtt 4260
gtcccacctg actccttgaa agtctttctgg acatgggaag ctatcgtatc gtatcgctcg 4320
ggcgaacatg atgtgtgtgt cactctcgag tgagcaggcc accgaaggct gacttgactg 4380
actccagcaa ccaacaaacg agccagtcat ttttcacccc gggttttttgt cccaaaaacac 4440
ttttccacca ccgtcaagcc tcaagcaaaa ccaaaacgct acgtaacgcc catcaacacc 4500
atgaaatcga gcagctagtt gtgcctgcta ctggcccccc agtgccctgt accgcccgtt 4560

ctctcactc gaca 4574